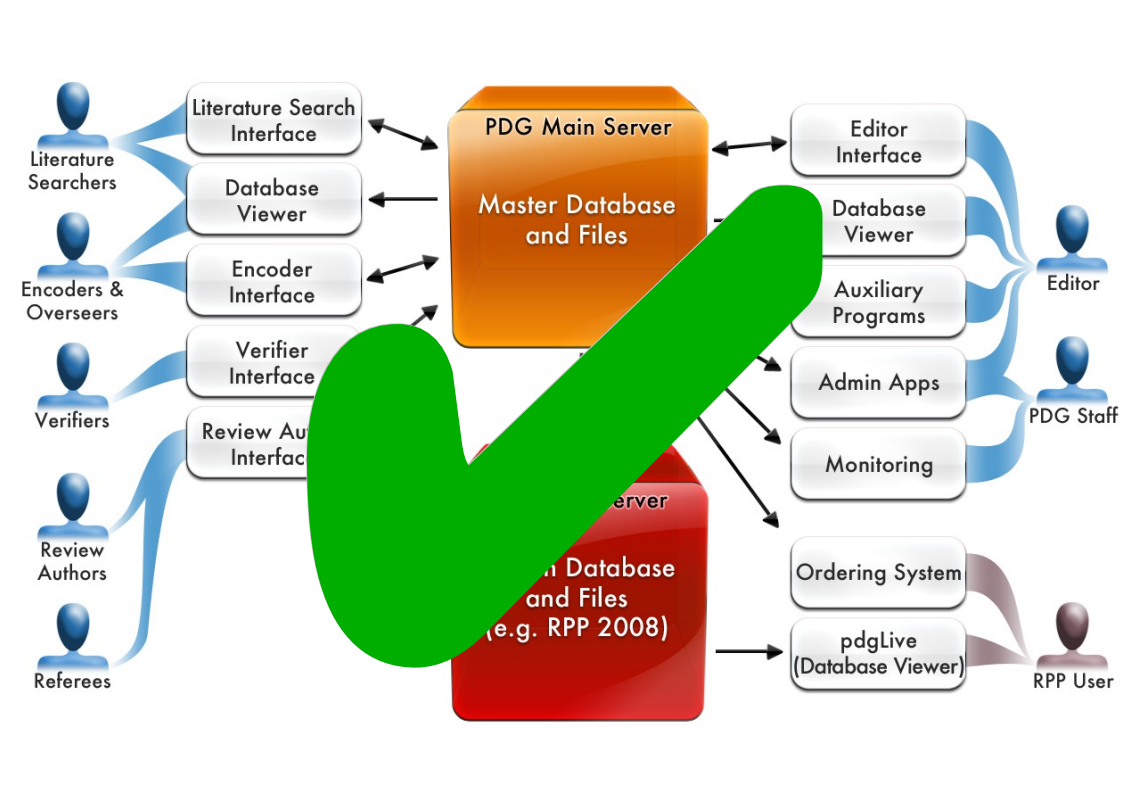
PDG Computing: Status and Future Directions

Juerg Beringer
Lawrence Berkeley National Laboratory

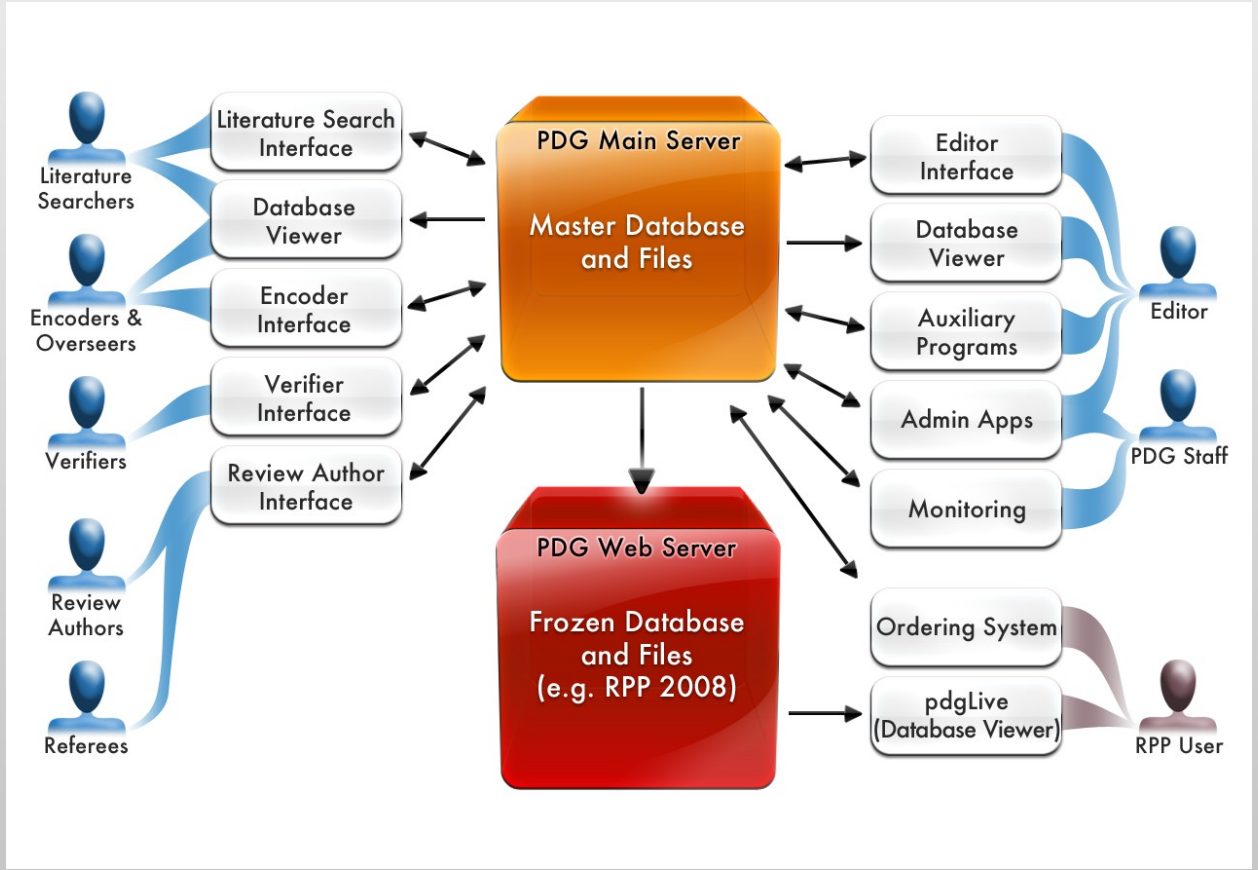
- **Summary of PDG Computing Upgrade**
- **(Your) Life with the new Computing System**
- **Discussion of future directions**

Summary of PDG Computing Upgrade



Goals and Scope

- **Developed and proposed plan for comprehensive upgrade of PDG computing infrastructure in 2007**
  - Primary goal: ensure PDG can continue to function well
  - Support all areas of PDG work, provide platform for future extensions
  - Set of web applications plus command-line production tools

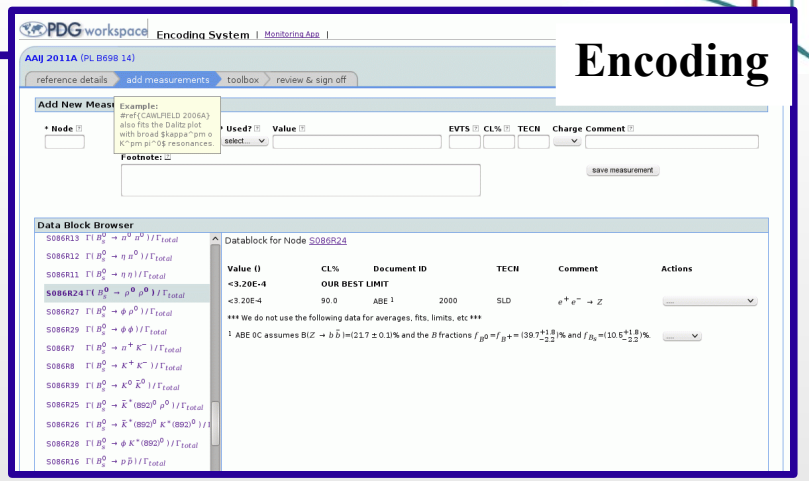


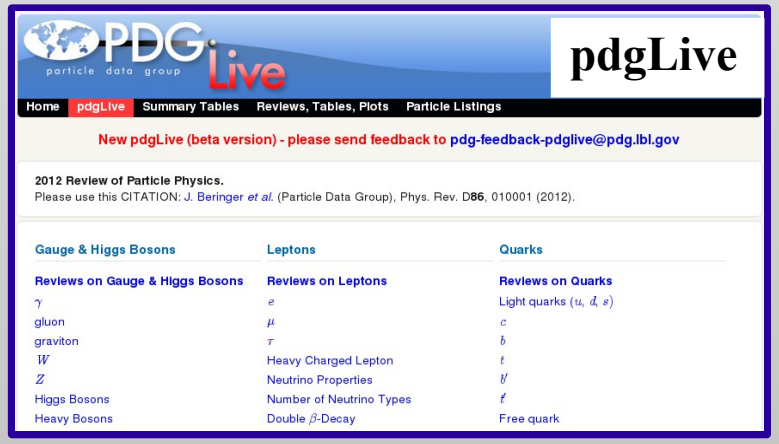
Brief History

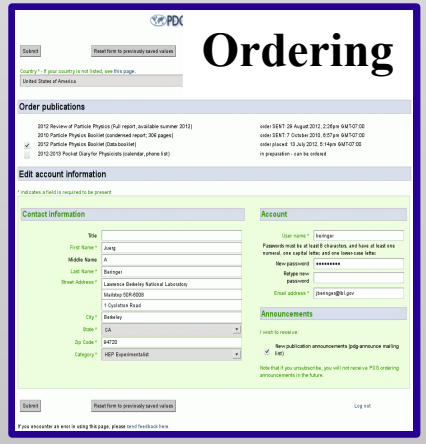
- **Project funded following DOE review of PDG in 2008**
  - 2 FTE for 3 years for development
  - 0.5 FTE (ongoing) for maintenance (starting in FY11)
- **Backbone of new system deployed into production in 2010**
- **DOE review of computing upgrade project on 9/17/2010**
  - "... upgrade is proceeding on schedule and within cost ..."
- **Very positive feedback at PDG meetings in Fall 2010**
- **Presentation/publication at CHEP'2010 (Taiwan)**
- **Had hoped to complete project early (in August 2011)**
  - Unexpected reduction in developer availability from May 2011 onwards meant we had to spread duration over full 3 years
  - Limited use of PdgWorkspace for 2012 edition
- **Project completed on schedule and budget in December 2011**

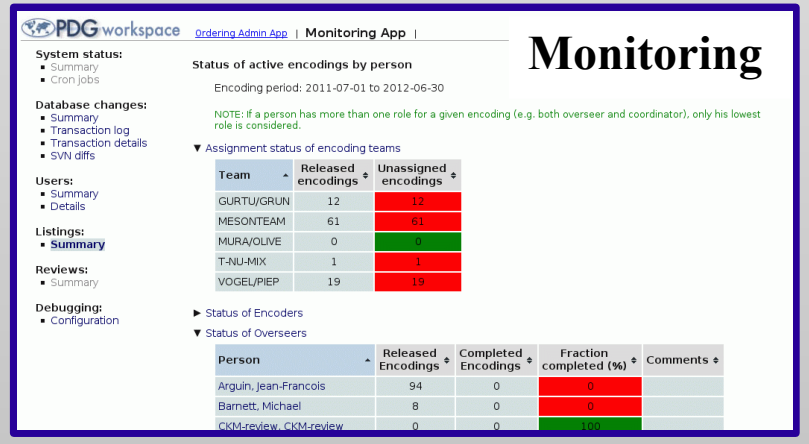
Results

- **All primary / high-priority web applications plus many secondary ones completed**
  - Very good progress on further features
- **Achieved much more regarding database modernization than thought possible**
- **Modern, extendable, maintainable and well-documented system**
- **Will only discuss subset here**



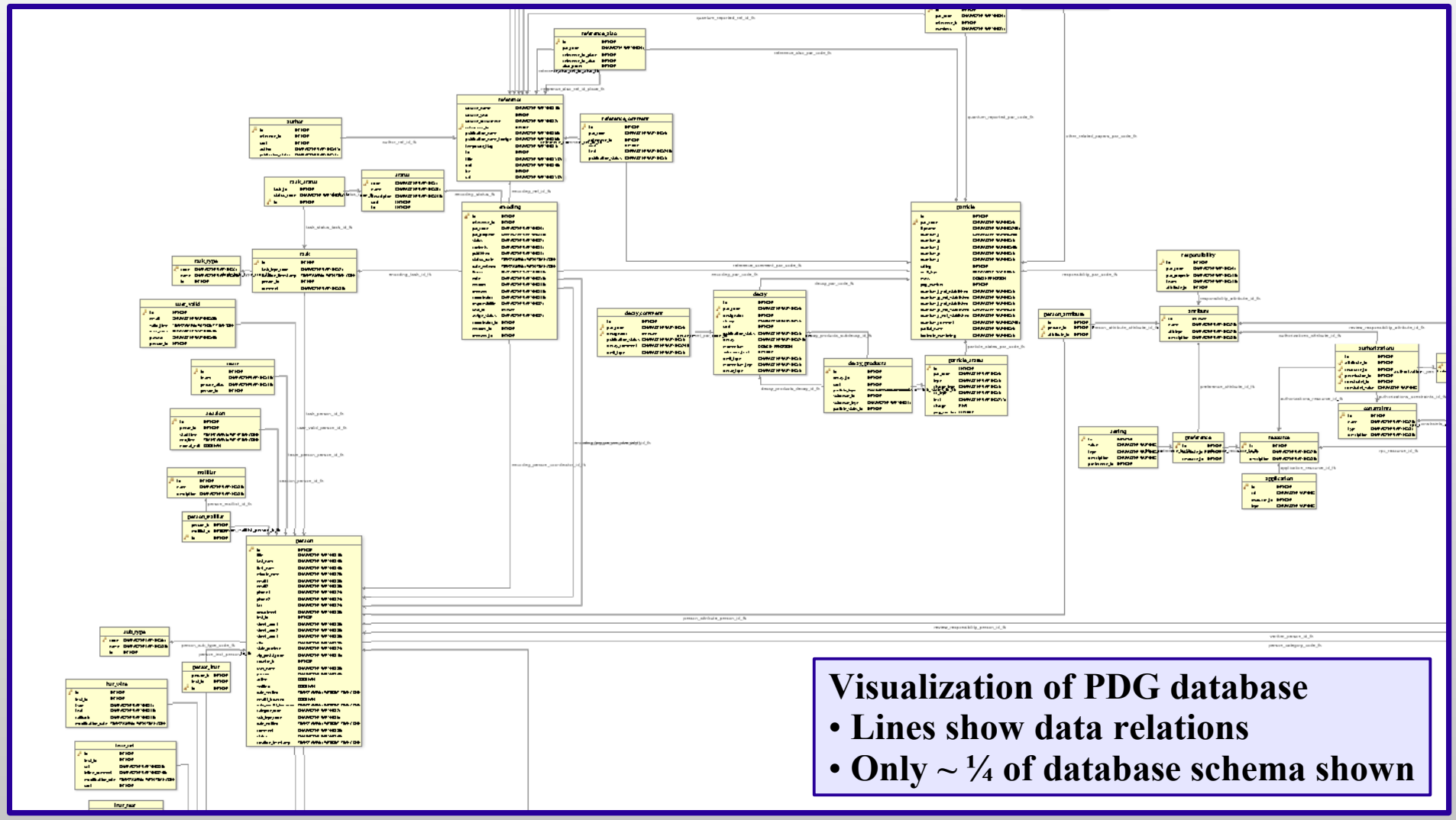






Modern PDG Database

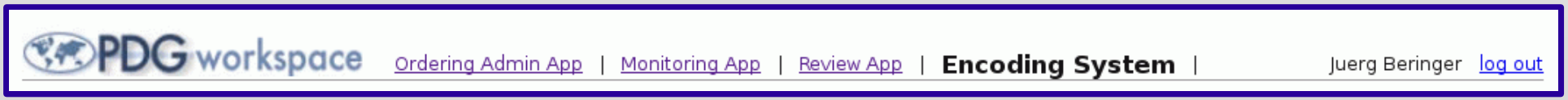
- **Large fraction of effort went into modernizing PDG database**
  - One of the **most valuable results of the project** in the long term



Visualization of PDG database
- Lines show data relations
- Only ~ ¼ of database schema shown

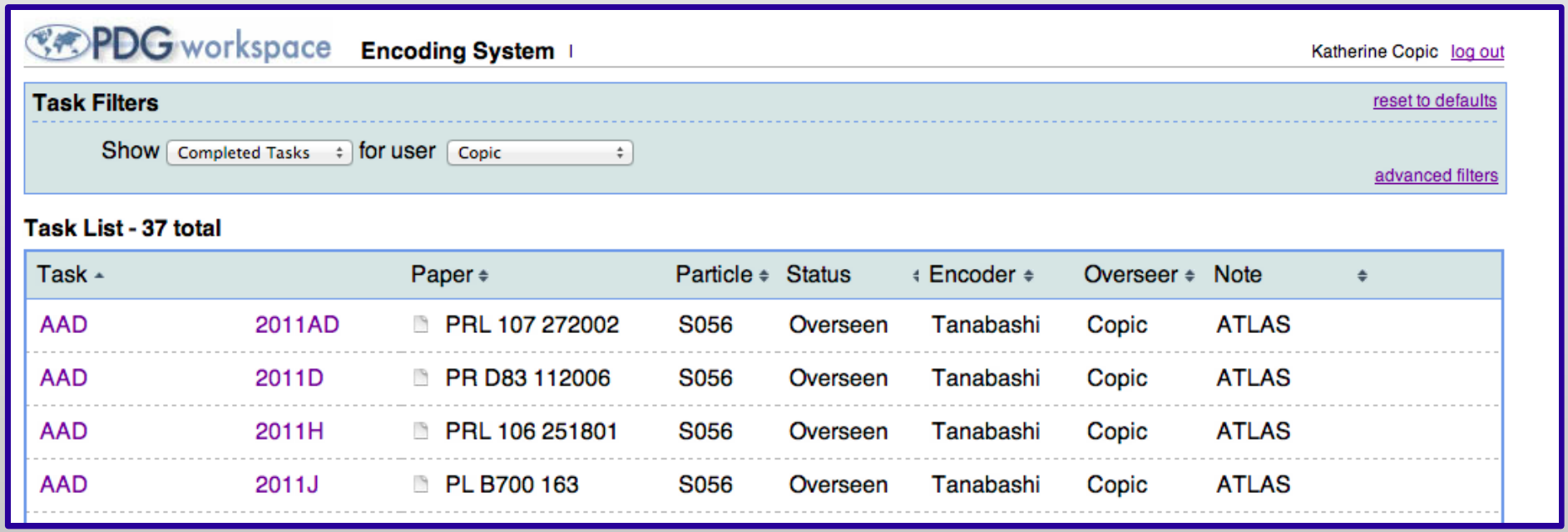
PDG Workspace

- **Framework used by all web applications for PDG collaborators**
  - This is <u>your</u> starting point for working with the new system
- **When you log in, you'll see a set of applications tailored to the work you do for PDG**



Encoding System

- **Tool for updating the Listings**
- **Successfully used by Kathy Copic for all her 2012 encodings**
  - Even being new to PDG, she **found system very easy to use**
- **Gradually invite encoders/overseers to start using new system**
  - Kathy Copic is contact person in case of problems



PDG workspace **Encoding System** | Katherine Copic log out

Task Filters reset to defaults

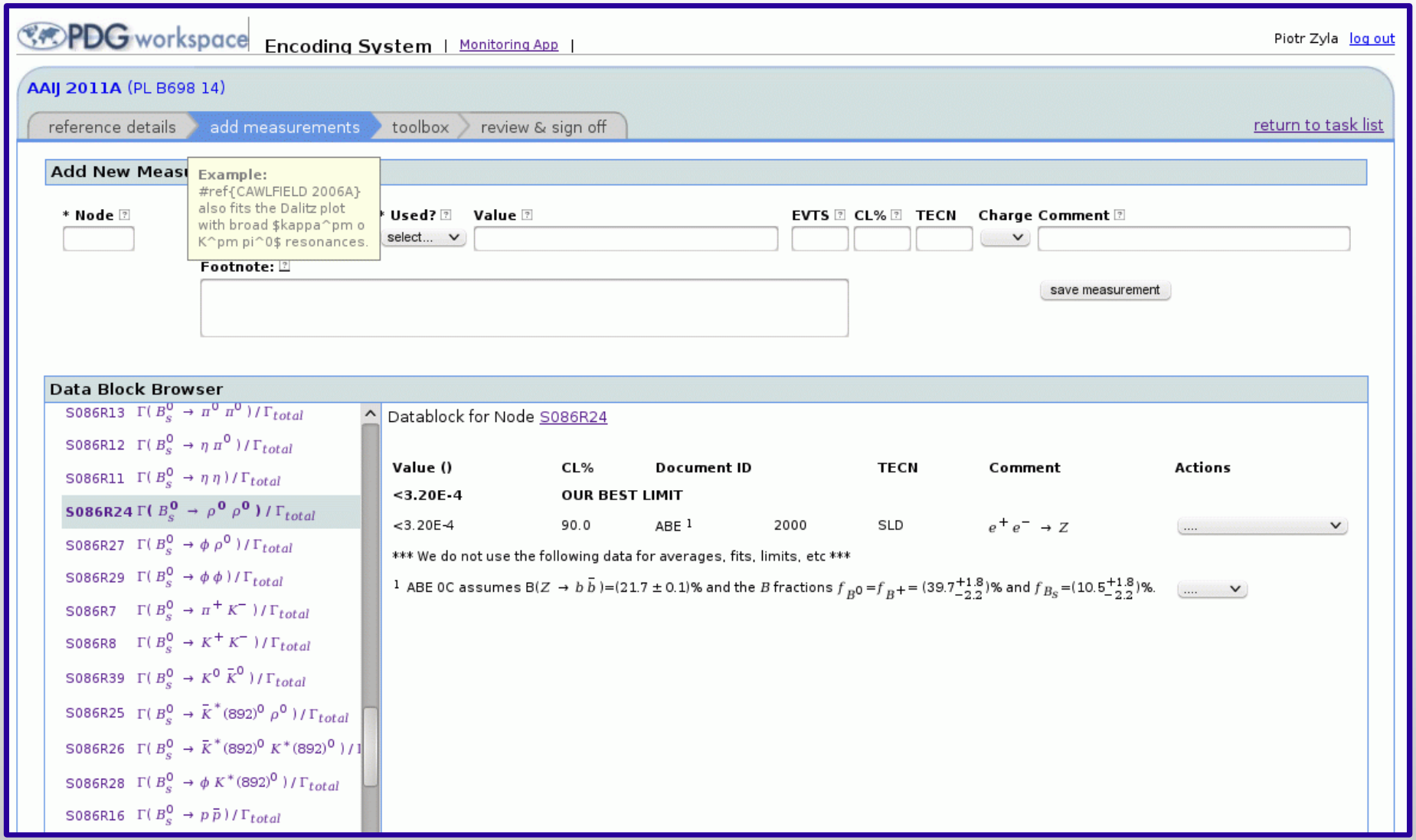
Show Completed Tasks for user Copic

advanced filters

Task List - 37 total

Task		Paper	Particle	Status	Encoder	Overseer	Note
AAD	2011AD	PRL 107 272002	S056	Overseen	Tanabashi	Copic	ATLAS
AAD	2011D	PR D83 112006	S056	Overseen	Tanabashi	Copic	ATLAS
AAD	2011H	PRL 106 251801	S056	Overseen	Tanabashi	Copic	ATLAS
AAD	2011J	PL B700 163	S056	Overseen	Tanabashi	Copic	ATLAS

Data Entry Page



Review Authoring System

- **Tool for authoring and refereeing review articles**
- **Supports 3 ways for authors to write reviews**
  - **Traditional**: check out tar file of sources, run TeX locally, return revised source files to overseer / editor, who updates them in system
  - **Using subversion (SVN)**: check out source files via SVN, commit changes when ready, run TeX either locally or get PDF file from review authoring system
  - **Online**: edit source files in your browser, no local TeX needed
- **Other features include**
  - Keeps track of review responsibilities, status and deadlines
  - History of changes and “diff” between versions in various forms
  - Links between reviews and related items (reviews, data blocks) in order to improve consistency across RPP
  - Integrated into overall progress monitoring
- **Will be used starting with 2013 review updates**

Review Task List

PDG workspace | Ordering Admin App | Monitoring App | **Review App** | Encoding System | Juerg Beringer log out

Review Filters

Show my reviews

Found 9 reviews

Title	Latest version	Finish draft by	Finish refereeing by	Status
Accelerator Physics of Colliders	PDF	2013-08-31	2013-09-30	Published
Free Quark Searches	PDF	2013-08-31	2013-09-30	Published
High-Energy Collider Parameters	PDF	2013-08-31	2013-09-30	Published
Monte Carlo Techniques	PDF	2013-08-31	2013-09-30	Published
Probability	PDF	2013-08-31	2013-09-30	Published
Statistics	PDF	2013-08-31	2013-09-30	Published
tau Branching Fractions	PDF	2013-08-31	2013-09-30	Published
tau-Lepton Decay Parameters	PDF	2013-08-31	2013-09-30	Published
The Top Quark	PDF	2013-08-31	2013-09-30	Published

PDG workspace | Ordering Admin App | Monitoring App | **Review App** | Encoding System | Juerg Beringer log out

return to review list

Review: High-Energy Collider Parameters

Details | Get source files | Edit online | History | Change review status

Monitoring: from Overall Progress ...

- **Used for 2012 edition and found to be very helpful**

PDG workspace | Ordering Admin App | **Monitoring App** | Juerg Beringer log out

System status:
- Summary
- Cron jobs

Database changes:
- Summary
- Transaction log
- Transaction details
- SVN diffs

Users:
- Summary
- Details

Listings:
- Summary

Reviews:
- Summary

Debugging:
- Configuration

Status of active encodings by person

Encoding period: 2011-07-01 to 2012-06-30

▶ Assignment status of encoding teams

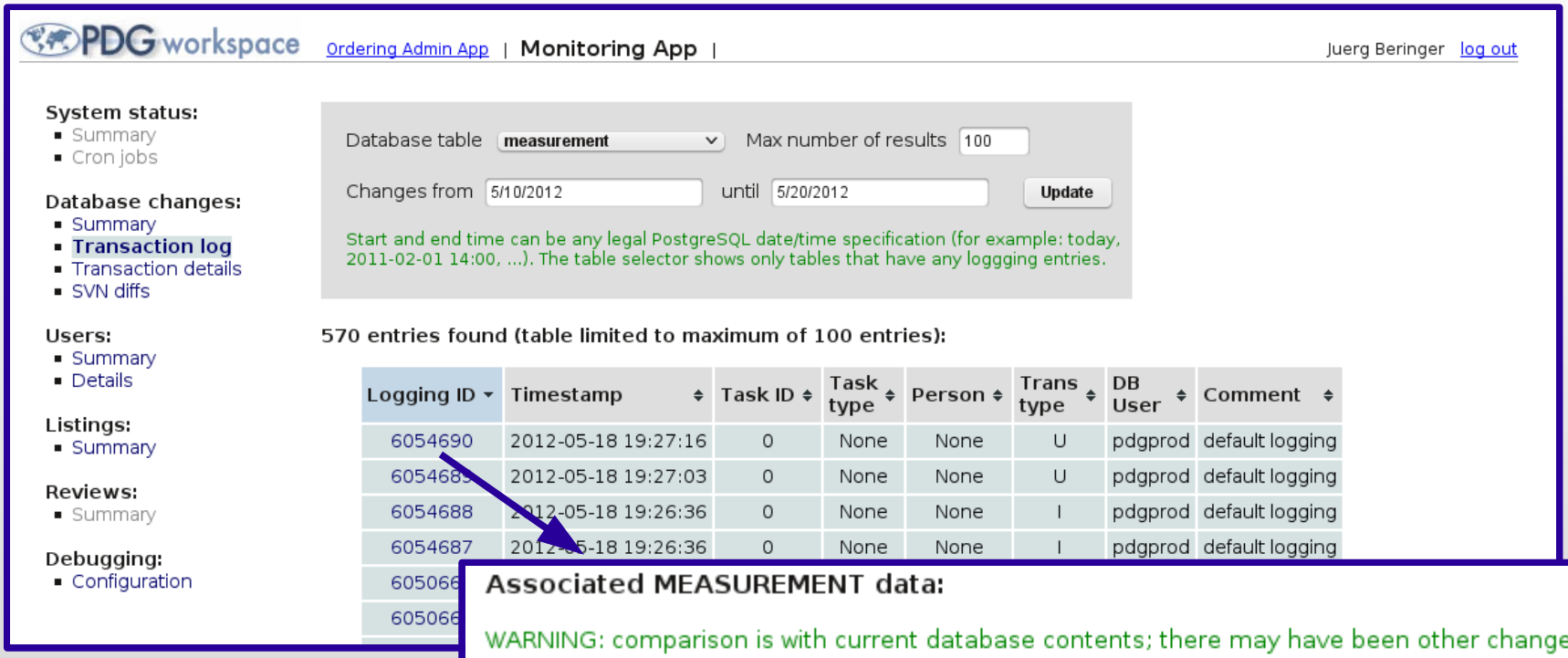
▶ Status of Encoders

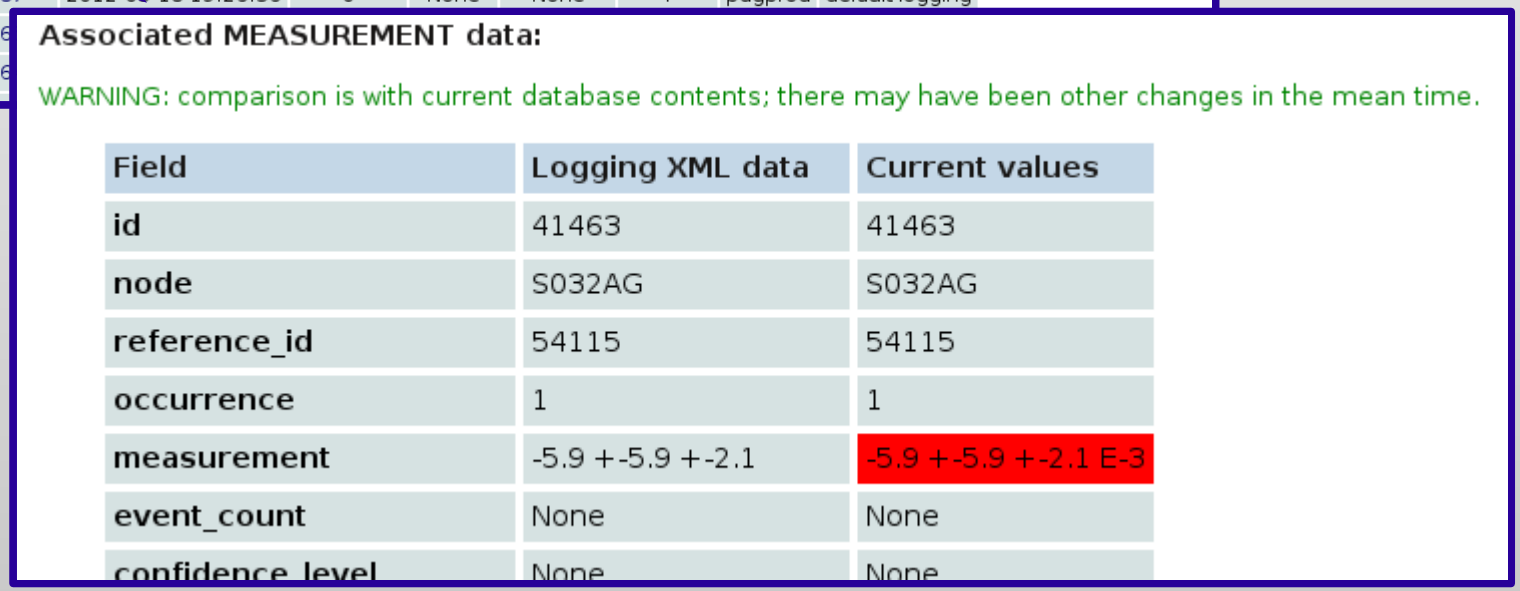
▼ Status of Overseers

Person	Released Encodings	Completed Encodings	Fraction completed (%)	Comments
Arguin, Jean-Francois	89	89	100	
Barnett, Michael	14	14	100	
Copic, Katherine	37	37	100	
Doser, Michael	86	86	100	
Groom, Don	7	7	100	
Gurtu, Atul	15	15	100	
Hikasa, Ken-ichi	55	55	100	
Lin, Cheng-Ju S	58	58	100	
M\"onig, Klaus	3	3	100	
Milstead, David	2	2	100	
Raffelt, Georg	20	20	100	
Terning, John	8	8	100	
Weiglein, Georg	97	97	100	
Wohl, Charles	82	82	100	
Yao, Wei-Ming	77	77	100	

▶ Status of Coordinators

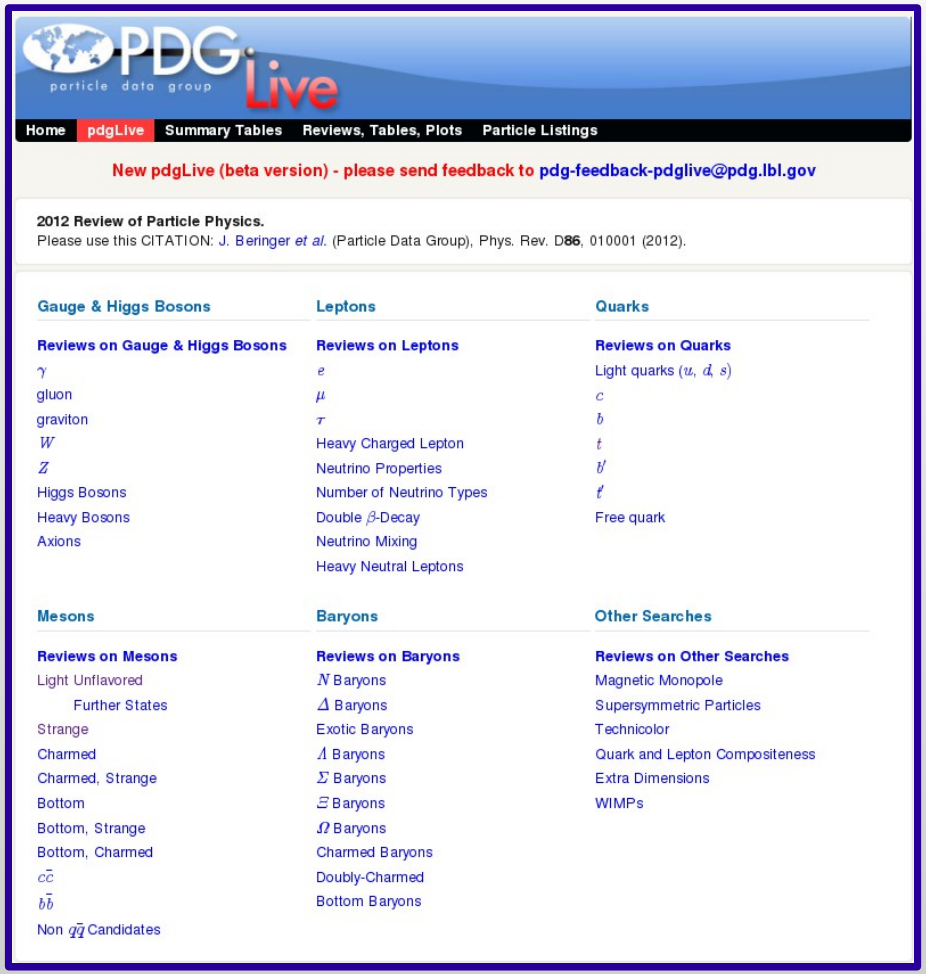
… to Detailed Database Changes





pdgLive

- **Same look-and-feel as "old" pdgLive**
- **Many improvements and new features**
  - See next talk for details
- **Deployed as "beta version" for RPP 2012, in parallel with old pdgLive**
  - Wanted to be cautious, since maintenance person went on maternity leave two weeks before deployment
  - Will become primary version of pdgLive after fixing a few known issues



Ordering System

- **Online since May 2011**
- **Will replace ordering via CERN library a few weeks after CERN ships 2012 booklets and diaries**
  - Full statistics & contact info in one place (so far: split LBNL & CERN)
  - CERN will still ship orders for Europe, Africa, ...
- **Users can subscribe to PDG mailing list(s)**
- **For PDG collaborators, same login as for PdgWorkspace**

(Your) Life with the New Computing System

The Plan

- **Want to gradually get collaborators started with new system**
  - Allows to **provide very good support** and fix any problems quickly, if needed
  - When the time comes, **we'll contact you** to start using new system
  - Let us know if you'd like to be among the first users
- **Start inviting encoders, overseers and literature searchers in December**
  - Use new system for 2013 encodings
- **Start inviting review authors in Spring/Summer 2013**
  - In time for next review update

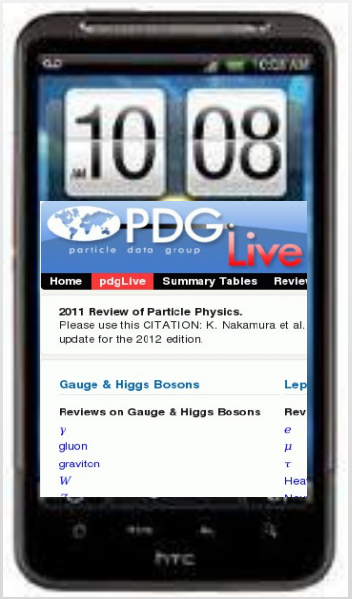
How to Get Started

- **Once it's time to start using the new system**
  - We will setup an account for you (unless you've used the ordering system and already created an account there)
  - We will enable this account to use the tools you need (literature search, encoding system, review author system, ...)
- **<u>After that</u>, you can log into PDG Workspace and start working**
  - https://pdgprod.lbl.gov/PdgWorkspace/
- **If you need help**
  - Documentation (TWiki): http://pdgprod.lbl.gov/twiki/bin/view/Pdg/
  - Each application has a responsible contact person – see Twiki
  - Contact your overseer, Piotr, or me
  - **Please let us know about any problems so that we can fix them!**

Discussion of Future Directions

New Possibilities

- **New PDG computing system provides solid basis for implementing new features**
  - Realizing substantial new features still requires a development effort
- **Many ideas / requests for new features from PDG**
  - **Smartphone app** – see later talk by Michael
  - **New features for pdgLive** – see next talk
  - **Data entry by collaborations?** - next slide
  - **Download of PDG data in machine-readable format**
    - For individual data blocks (XML, via pdgLive)
    - All PDG data (in a SQLite database)
  - Extending PDG computing platform to support averaging groups?
  - Programmatic access to PDG database outside of PDG collaboration?
  - ...



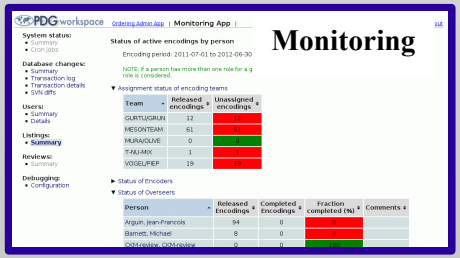
Data Entry by Collaborations?

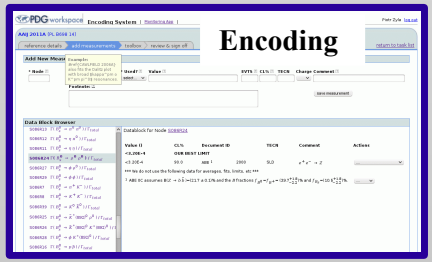
- **Should we allow the community to enter new results directly into the PDG database?**
  - In particular large collaborations such as LHC experiments
  - A result is best known/understood by its author
  - **PDG would still validate all entries to be published in RPP**
    - Could start with a draft entry rather than just the article
    - Easier to include latest results
  - **In principle, pdgLive, could show user-contributed data in addition to data vetted by PDG**
    - **Users choose whether they want to see user-contributed entries**
    - User-contributed entries could appear online immediately
    - User-contributed data might include unpublished (public) data
    - Fits and averages including user-contributed data possible
  - Should be able to **minimize abuse** by requiring authentication and displaying the identity of contributiors

Is this something worth pursuing?

Conclusions

- **PDG Computing Upgrade project successfully completed on schedule and within budget in December 2011**
  - Backbone of new system used in production since summer 2010
  - Encoding interface and monitoring used for producing RPP 2012
- **As a result of this project, PDG now has a solid computing platform on which we can build**
  - Modern, extendable, maintainable and well-documented
- **Gradually inviting collaborators to start using new system**
  - Will allow us to provide very good support and fix problems quickly
- **Discussion about future direction**





Backup Slides

Intrinsically Challenging Project

- **Convert ancient legacy software into modern system without disrupting ongoing PDG work**
  - Gradually transform O(100) disconnected data tables into modern relational database with integrity constraints and change tracking
- **Complex distributed workflow with hundreds of users**
- **Correctness of data is paramount**
  - New system produces “identical” manuscript of 1400+ page book
- **Display of math on the web**
- **PDG schedule leaves only few deployment opportunities**
  - Gradual deployment, starting in August 2010

Our Approach

- **Focus and prioritize**
  - Identified technically challenging areas: **database, math display**
  - **Focus on cornerstone of new system**: modern PDG database
  - **Complete primary and difficult applications / middleware first**
  - Secondary applications lower priority than publication-critical ones
    - Remaining secondary web applications small and easy to complete with available resources
- **Avoid scope creep and keep to original goals**
  - Sometimes hard, since new system offers so many exciting opportunities for extension beyond the project's scope
- **Strong involvement from PDG group**
  - Orin Dahl, Piotr Zyla, JB
  - Many others helped with testing

Documentation

- **Excellent documentation emphasized from beginning**
  - Manuals
  - TWiki
  - Inline code documentation, man pages, help

Public, for PDG collaborators:

TWiki > Pdg Web > WebHome (2011-12-15, Main.beringer) Edit Attach

Welcome to the TWiki of the Particle Data Group

This TWiki provides **information for PDG collaborators** about the new PDG computing system. Other information is available:

- The public PDG web pages can be found at http://pdg.lbl.gov.
- For general information about working in the PDG collaboration please refer to the Encoder Tools section of the PDG web page.
- Information for PDG Computing developers can be found in the Computing TWiki (protected).

How to get Started with the New PDG Computing System

The new PDG Computing System supports all activities of the PDG collaboration through a set of web-based tools. Most of these tools are accessed through a single portal called PdgWorkspace. The URL of this portal is:

http://pdgprod.lbl.gov/PdgWorkspace/

In order to access PdgWorkspace, you need a PDG web account with a username (typically your e-mail address) and a password. If you ever used the new PDG ordering system, you already have a PDG web account. Otherwise a new one will be created for you. Your PDG web account needs to be authorized and associated with your roles in the PDG collaboration before you can use it to access PdgWorkspace.

The new PDG computing system and the use of PdgWorkspace are currently being phased in. We will contact each collaborator in due course with information about their PDG web account. However, if you do want to jump right in as soon as possible, you may request the creation and authorization of your PDG web account by contacting Piotr Zyla.

Brief Description of the different tools in PdgWorkspace

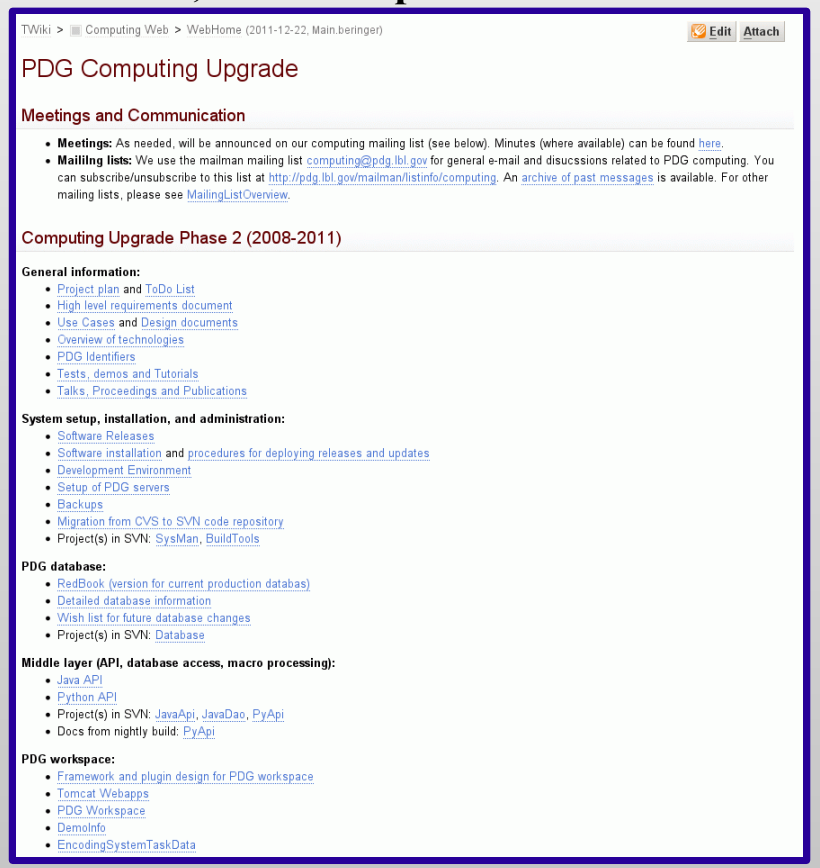
When you log into PdgWorkspace, the system knows your role and responsibilities in PDG, and will give you access to the corresponding tools. These tools may include:

- **Encoding Tool**: The encoding tool supports the full encoding cycle for entering new results into the Listings. It is used by Literature Searchers, Encoder, Overseers, and Coordinators.
- **Review Author Tool**: The review author tool supports the authoring of review articles. It is used by review authors, review coordinators and referees *(deployed starting summer 2012)*.
- **Monitoring Tool**: The monitoring tool is a comprehensive tool for monitoring the status of PDG work as well as the new computing system itself.
- **Administrative tools**: Various administrative tools support different tasks such as managing the ordering system, creating user accounts, etc.

Frequently Asked Questions (and Answers)

- **I didn't find what I need - where can I get help?** Depending on what you're looking for, please contact one of the persons listed below:
  - Questions about how to encode something for the Listings: Piotr Zyla
  - Technical questions about PdgWorkspace: Sarah Poon
  - Technical questions about the PDG ordering system: Sarah Poon
  - Special requests regarding ordering of PDG products: pdgrequest@lbl.gov
  - General questions, or if you are unsure whom to ask: Juerg Beringer

Protected, for developers:



TWiki > Computing Web > WebHome (2011-12-22, Main.beringer) Edit Attach

PDG Computing Upgrade

Meetings and Communication

- **Meetings:** As needed, will be announced on our computing mailing list (see below). Minutes (where available) can be found here.
- **Mailing lists:** We use the mailman mailing list computing@pdg.lbl.gov for general e-mail and disucssions related to PDG computing. You can subscribe/unsubscribe to this list at http://pdg.lbl.gov/mailman/listinfo/computing. An archive of past messages is available. For other mailing lists, please see MailingListOverview.

Computing Upgrade Phase 2 (2008-2011)

General information:

- Project plan and ToDo List
- High level requirements document
- Use Cases and Design documents
- Overview of technologies
- PDG Identifiers
- Tests, demos and Tutorials
- Talks, Proceedings and Publications

System setup, installation, and administration:

- Software Releases
- Software installation and procedures for deploying releases and updates
- Development Environment
- Setup of PDG servers
- Backups
- Migration from CVS to SVN code repository
- Project(s) in SVN: SysMan, BuildTools

PDG database:

- RedBook (version for current production databas)
- Detailed database information
- Wish list for future database changes
- Project(s) in SVN: Database

Middle layer (API, database access, macro processing):

- Java API
- Python API
- Project(s) in SVN: JavaApi, JavaDao, PyApi
- Docs from nightly build: PyApi

PDG workspace:

- Framework and plugin design for PDG workspace
- Tomcat Webapps
- PDG Workspace
- DemoInfo
- EncodingSystemTaskData

Knowledge Transfer

- **Hands-on involvement of PDG personnel ensures excellent knowledge of new system from start**
- **Schema migration tutorial/exercise on 12/21/2011**
  - Using a copy of the new system locally installed on our laptops, successfully modified the database schema, updated all necessary software components, and recompiled everything
- **Sarah Poon retained as maintenance person**
  - On maternity leave since 6/13/2012
- **Other developers still accessible at LBNL if needed**

Project Team

- **Juerg Beringer (PDG physicist)**
  - Project leader, requirements, system architecture
- **Chuck McParland (computer scientist)**
  - Java API
- **Sarah Poon (computer systems engineer)**
  - Web design, user interfaces, JavaScript
- **David Robertson (computer systems engineer)**
  - Database, Python API, scripts
- **Orin Dahl (PDG physicist, retired)**
  - Legacy Fortran programs
- **Piotr Zyla (PDG editor)**
- **Contributions from Jacob Andreas, Cecilia Aragon, Keith Beattie, Igor Gaponenko, Keith Jackson, Kirill Lugovsky, Slava Lugovsky**

Each member of the team has many years of software development experience

Automated Testing

from PDG Nightly Builds <pdg-build@george.lbl.gov>
subject **PDG nightly build report**
to PDG Build list <pdg-build@george.lbl.gov>

```
pdg-job-runner.sh: Wed May 30 04:01:01 PDT 2012 starting...

  success pdgbuild-recreate.sh in 112 secs
  success python.sh in 4 secs
  success JavaApi in 39 secs
  success JavaDao in 24 secs
  success AuthApi in 25 secs
  success JavaModels in 23 secs
* failure PdgWorkspace (1) in 32 secs
  success OrderingApp in 29 secs

See http://pdg6.lbl.gov/cgi-bin/pdg-monitor.cgi for details
pdg-job-runner.sh: Wed May 30 04:05:50 PDT 2012 done.
```

```
[WARNING] /home/pdgbuild/pdgbuild-work/PdgWorkspace/src/main/java/pdg/workspace/helpermodels/EncodingTask.java:39: warning - @return tag has no arguments.
[WARNING] /home/pdgbuild/pdgbuild-work/PdgWorkspace/src/main/java/pdg/workspace/ext/BranchingRatioFormatter.java:58: warning - @param argument "s" is not a parameter name.
[INFO] [resources:resources {execution: default-resources}]
[INFO] Using 'UTF-8' encoding to copy filtered resources.
[INFO] Copying 0 resource
[INFO] [compiler:compile {execution: default-compile}]
[INFO] Compiling 39 source files to /home/pdgbuild/pdgbuild-work/PdgWorkspace/target/classes
[INFO] -------------------------------------------------------------
[ERROR] BUILD FAILURE
[INFO] -------------------------------------------------------------
[INFO] Compilation failure
/home/pdgbuild/pdgbuild-work/PdgWorkspace/src/main/java/pdg/workspace/ext/StringFormatter.java:[128,49] cannot find symbol
symbol  : method getDisplayText()
location: class pdg.macro.util.displayPrimitive.DisplayMeasurementTex


[INFO] ------------------------------------------------------------------------
[INFO] For more information, run Maven with the -e switch
[INFO] ------------------------------------------------------------------------
[INFO] Total time: 28 seconds
[INFO] Finished at: Wed May 30 04:05:20 PDT 2012
[INFO] Final Memory: 59M/981M
[INFO] ------------------------------------------------------------------------
building done.
```

Literature Search Tool